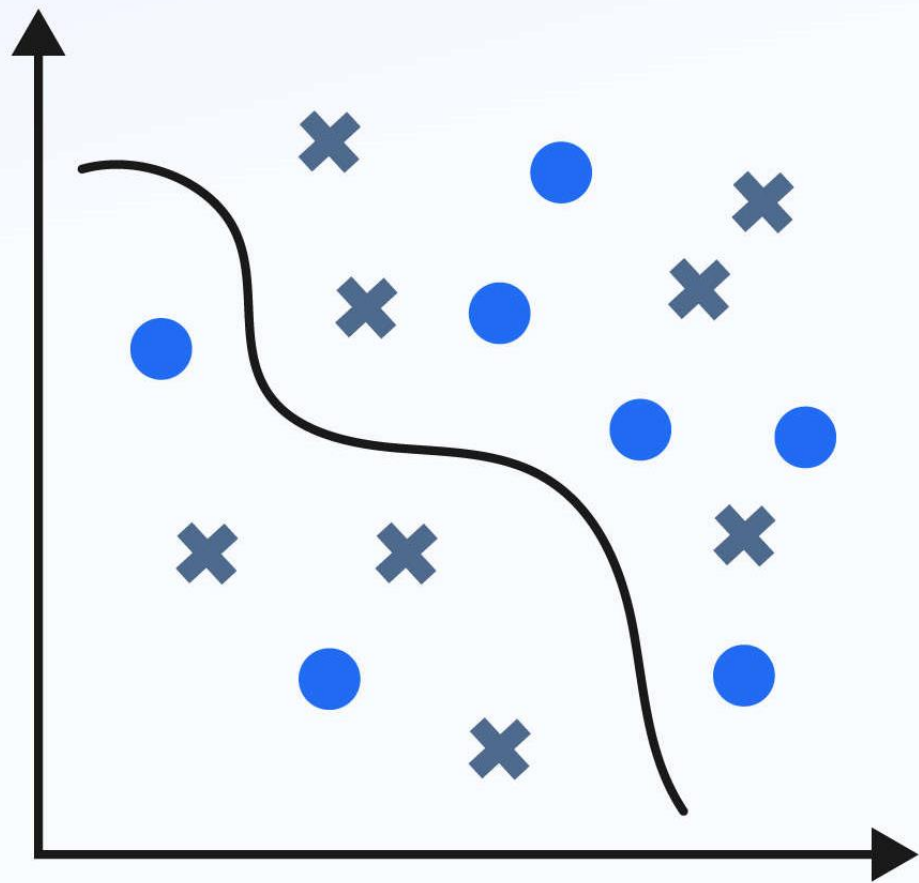


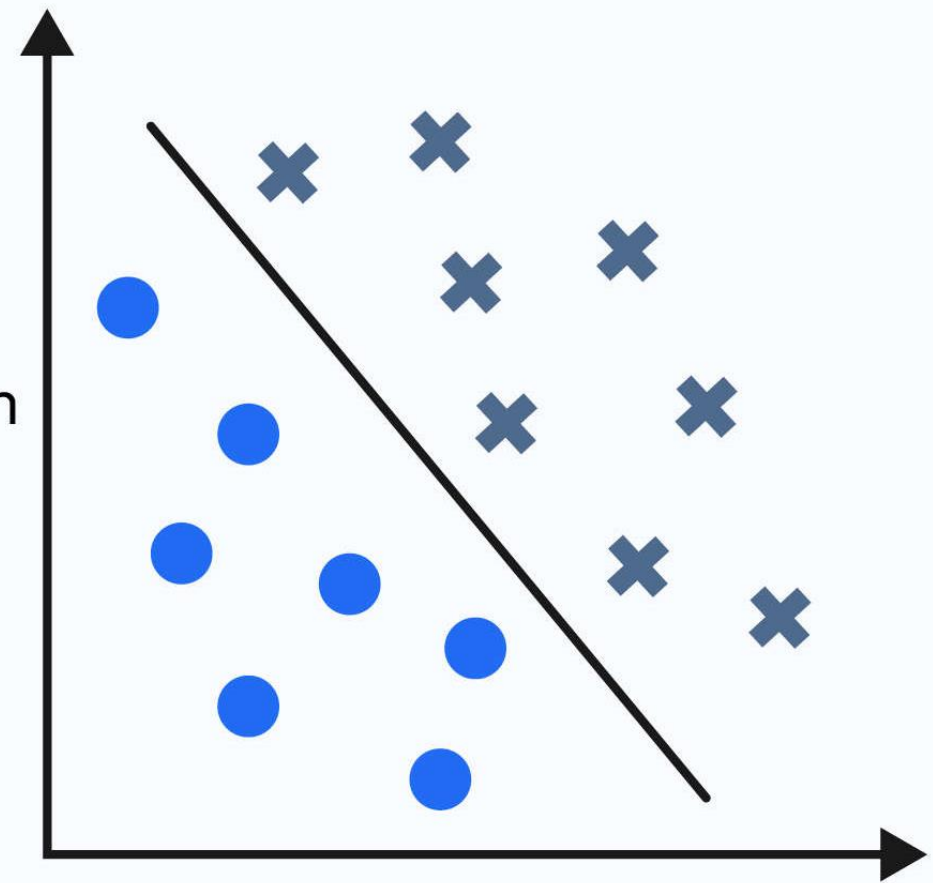
Feature Extraction & Analysis

Session-13



Raw Data

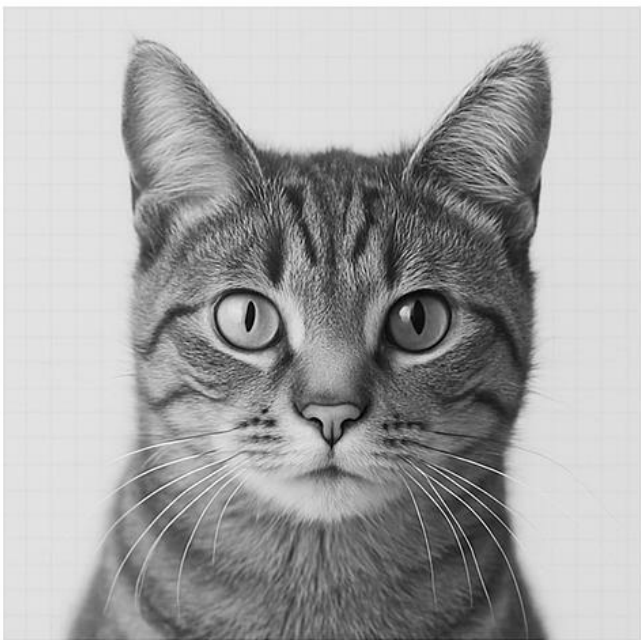
Feature
Extraction
→



Discriminative Features

Significance of Feature Extraction

Aspect	Description
Dimensionality Reduction	Reduces high-dimensional image data into compact, relevant feature sets.
Improved Accuracy	Enhances model performance by focusing on meaningful patterns and structures.
Robustness	Makes analysis resilient to noise, scaling, rotation, and lighting changes.
Machine Learning Readiness	Converts image data into numerical features usable by algorithms.
Speed and Efficiency	Decreases computational load and accelerates processing and training.
Application Flexibility	Enables a wide range of tasks like classification, detection, and segmentation.
Feature Types	Includes hand-crafted (e.g., SIFT, HOG) and deep learning-based (CNN) features.



**High-
Dimensional
Data**

**Feature
– Extraction →**



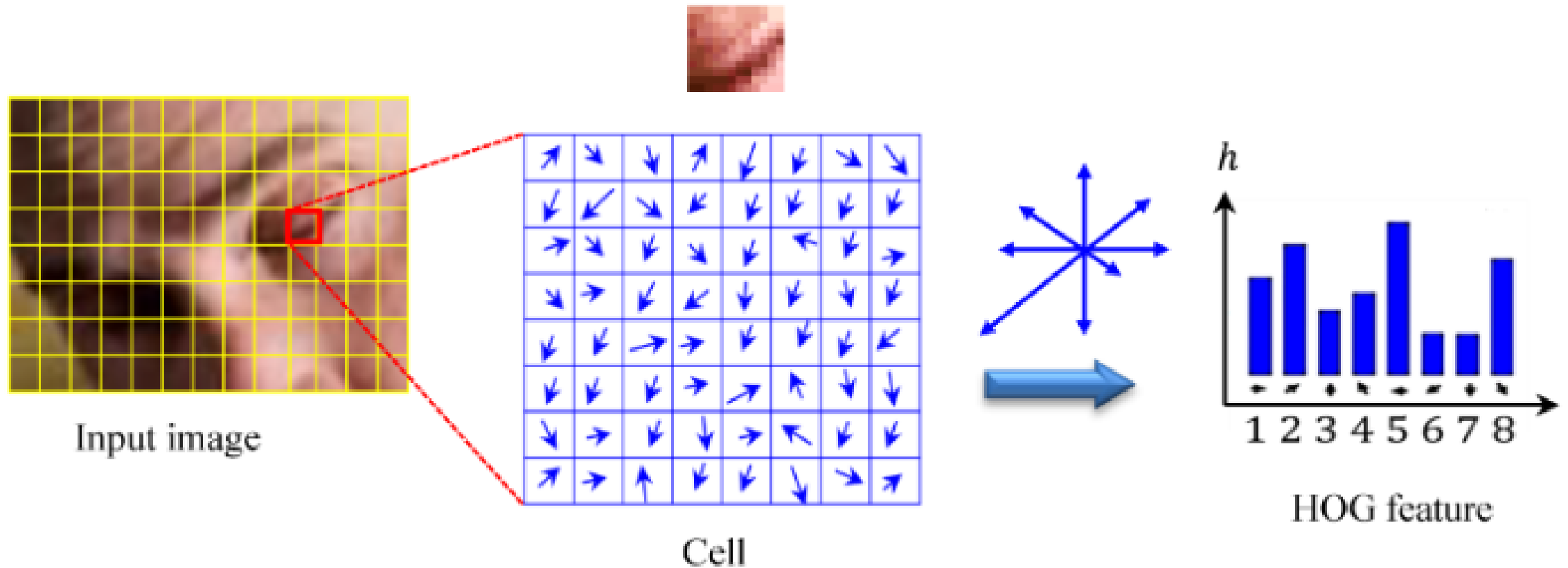
**Compact
Features**



HOG & SIFT Features

- The most critical factor for object detection is the gradient.
 - Both HOG and SIFT are feature descriptors based on the gradient orientation histogram in the localized portion of an image.
- As these two techniques are quite similar to each other, it can easily for beginners to get confused about these two concepts in computer vision and image processing.
 - To calculate the gradient, the grayscale image is used.
 - The illumination of the environment makes it difficult to provide key information. Therefore, we need to convert the colour image into a grayscale image to speed up feature extraction.

HOG Feature Calculation



HOG (Histogram of Oriented Gradients)

Histogram of Oriented Gradients, also known as HOG, is a feature descriptor that focuses on the structure or the shape of an object.

- **Steps to Calculate HOG Features**

1. Take the input image, and convert the colour image to Grayscale;
2. Crop the region of interest, and resize it into an image of 128x64 pixels ;
3. Use the Gamma correction method to standardize (normalize) the colour space of the input image.
4. To adjust the contrast of the image, reduce the impact of local shadows and illumination changes in the image, and at the same time suppress the interference of noise;
5. Divide the image into small cells (eg., 8*8 pixels/cell);
6. Count the gradient histogram of each cell (the number of different gradients) to form the descriptor of each cell;
7. Every several cells form a block (for example, 2*2 cells/block), and the feature descriptors of all cells in a block are concatenated to obtain the HOG feature descriptor of the block. For each block, a 9-point histogram is calculated with 9 bins and each bin has an angle range of 20 degrees.
8. Concatenate the HOG feature descriptors of all blocks in the image to get the HOG feature descriptor of the image (the target you want to detect). This is the final feature vector available for classification.

HOG Feature Calculation

- Calculate the gradient (including size and direction) of each pixel of the image, mainly to capture contour information and further weaken the interference of light;

$$Magnitude(\mu) = \sqrt{G_x^2 + G_y^2} \quad Angle(\theta) = |\tan^{-1}(G_y/G_x)|$$

Application

- HOG features are especially suitable for human detection in images.
- Under the conditions of coarse airspace sampling, fine direction sampling, and strong local optical normalization, as long as the pedestrian can generally maintain an upright posture, some subtle body movements can be allowed.
- These subtle movements can be ignored without affecting the detection result.



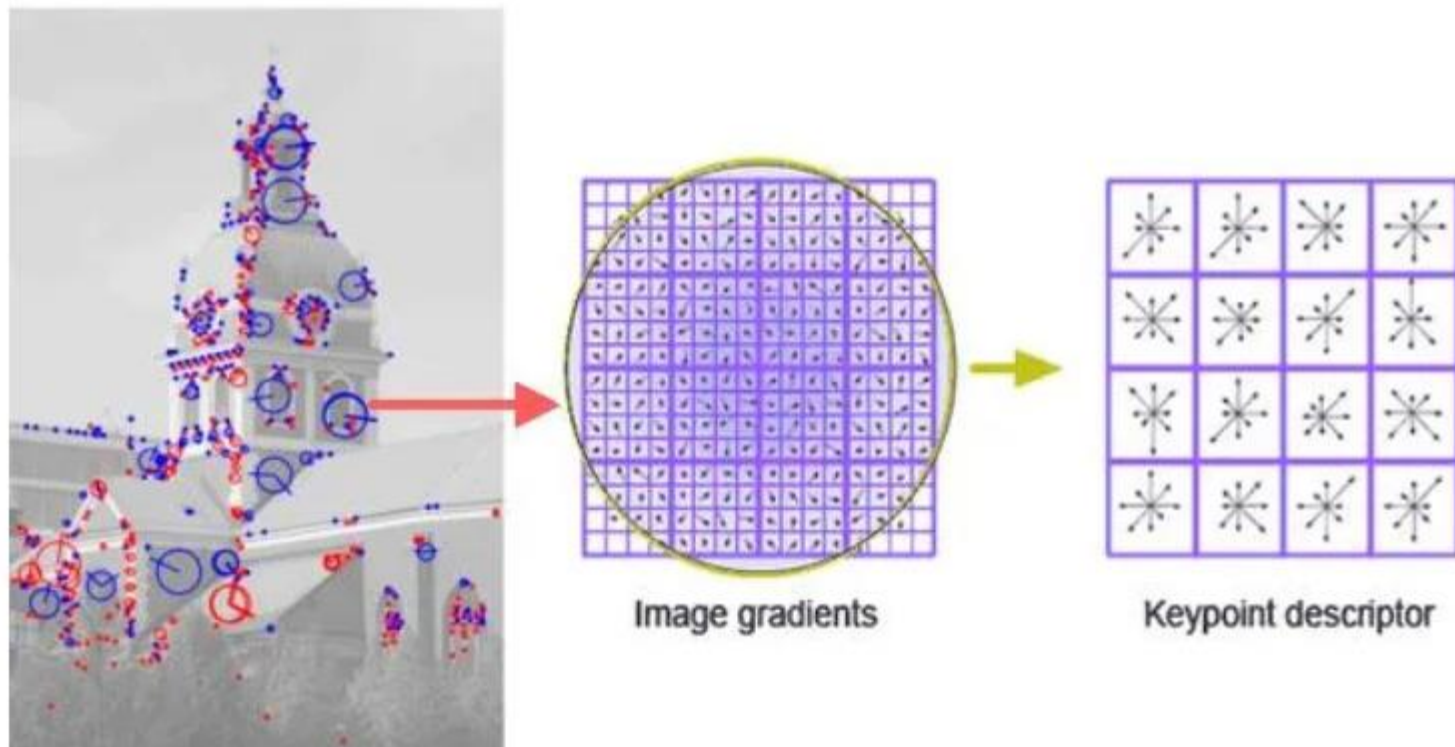
SIFT (Scale-Invariant Feature Transform)

SIFT stands for Scale-Invariant Feature Transform and is invariant to image scale and rotation.

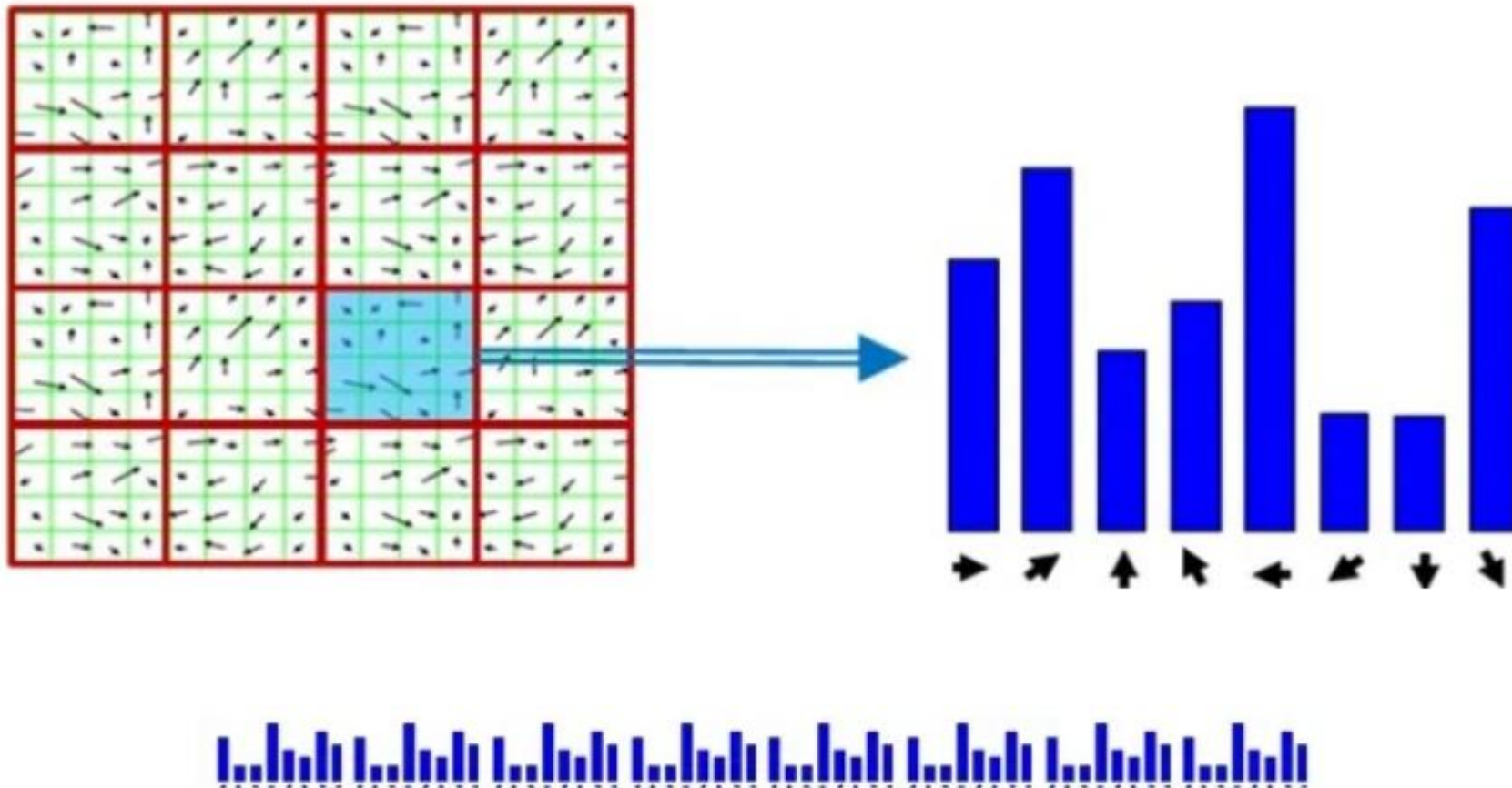
Find key points (feature points) in different scale-spaces, and calculate the direction of key points.

The key points found by SIFT are prominent points that are invariant to factors such as lighting, affine transformation, and noise, such as corner points, edge points, bright spots in dark areas, and dark points in bright areas.

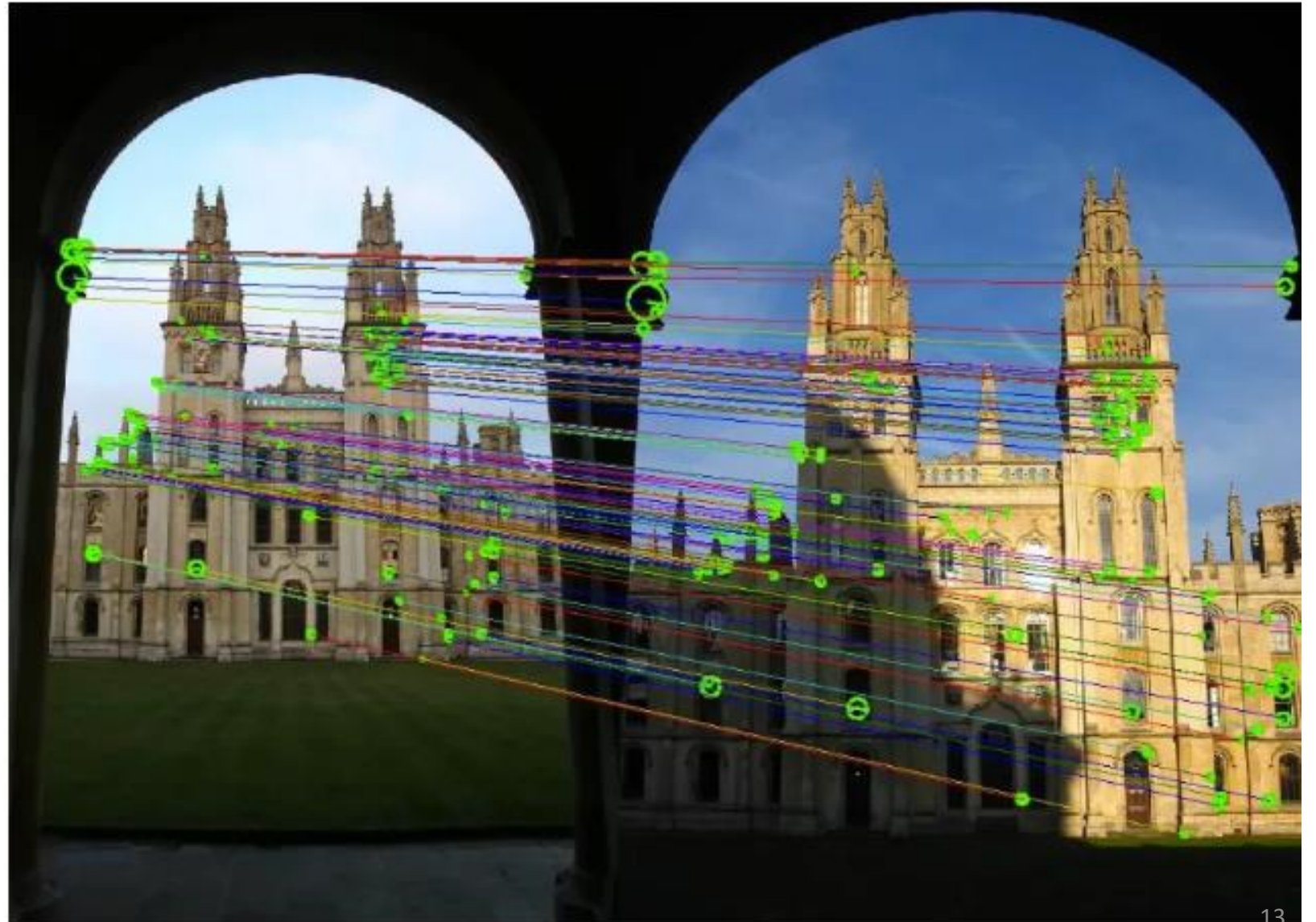
SIFT – Keypoint Localization



SIFT: Generation of Keypoint Descriptors



Feature matching and Keypoint Detection Using SIFT



Comparison of HOG and SIFT Method

- SIFT features are usually used together with interest points obtained using SIFT detectors. These points of interest are associated with a specific orientation and scale. Usually, the SIFT feature of the square area in an image is calculated after the corresponding direction and scale transformation are performed.
- The unit size of the HOG feature is small, so a certain spatial resolution can be preserved, and the normalization operation makes the feature insensitive to local contrast changes.
- **Comparing these 2 methods, SIFT is better for feature extraction of objects in complex environments; while HOG is good for feature extraction of rigid objects.**
- **Based on the experiment of natural scene classification experiment, it can be found that the accuracy of SIFT is higher than that of HOG. HOG's performance is better than SIFT, especially when detecting rigid objects.**

Comparison of HOG and SIFT Method

Aspect	HOG	SIFT
Purpose	Captures edge or gradient structure of objects	Detects and describes local features (keypoints)
Feature Type	Dense descriptor (whole image or grid regions)	Sparse descriptor (keypoints only)
Invariance	Partially invariant to scale and rotation	Fully invariant to scale, rotation, and partially to affine transforms
Computation	Faster and simpler	More computationally intensive
Descriptor Size	Fixed-length (e.g., 3780 for 64x128 window)	Typically 128-dimensional per keypoint
Use Case	Human detection, object recognition	Object recognition, image matching, stitching
Robustness to Illumination	Moderate	High
Gradient Use	Uses orientation histograms in localized regions	Uses orientation histograms around keypoints
Patent Status	Open-source and widely available	Originally patented (now expired)

Image Analysis Techniques

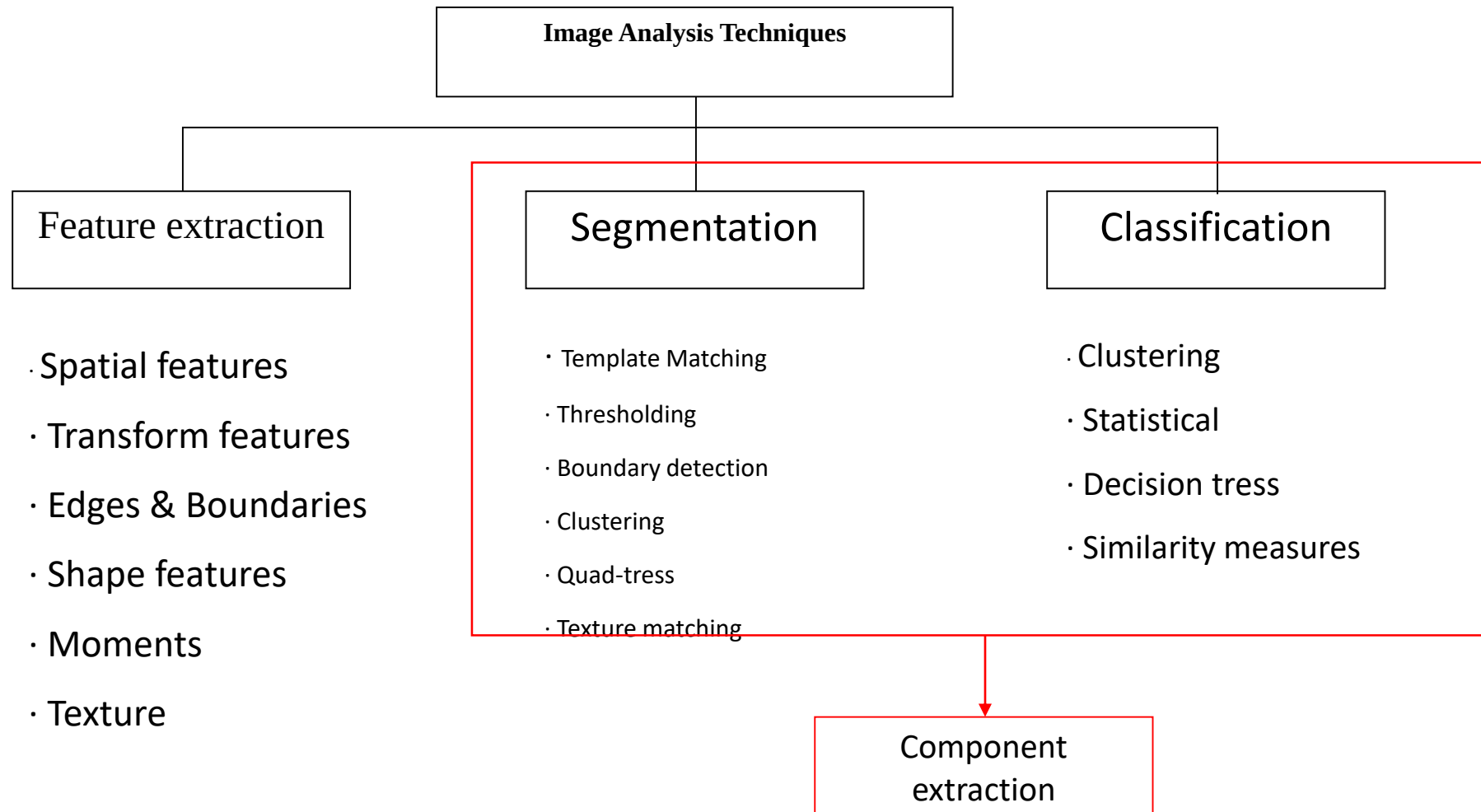
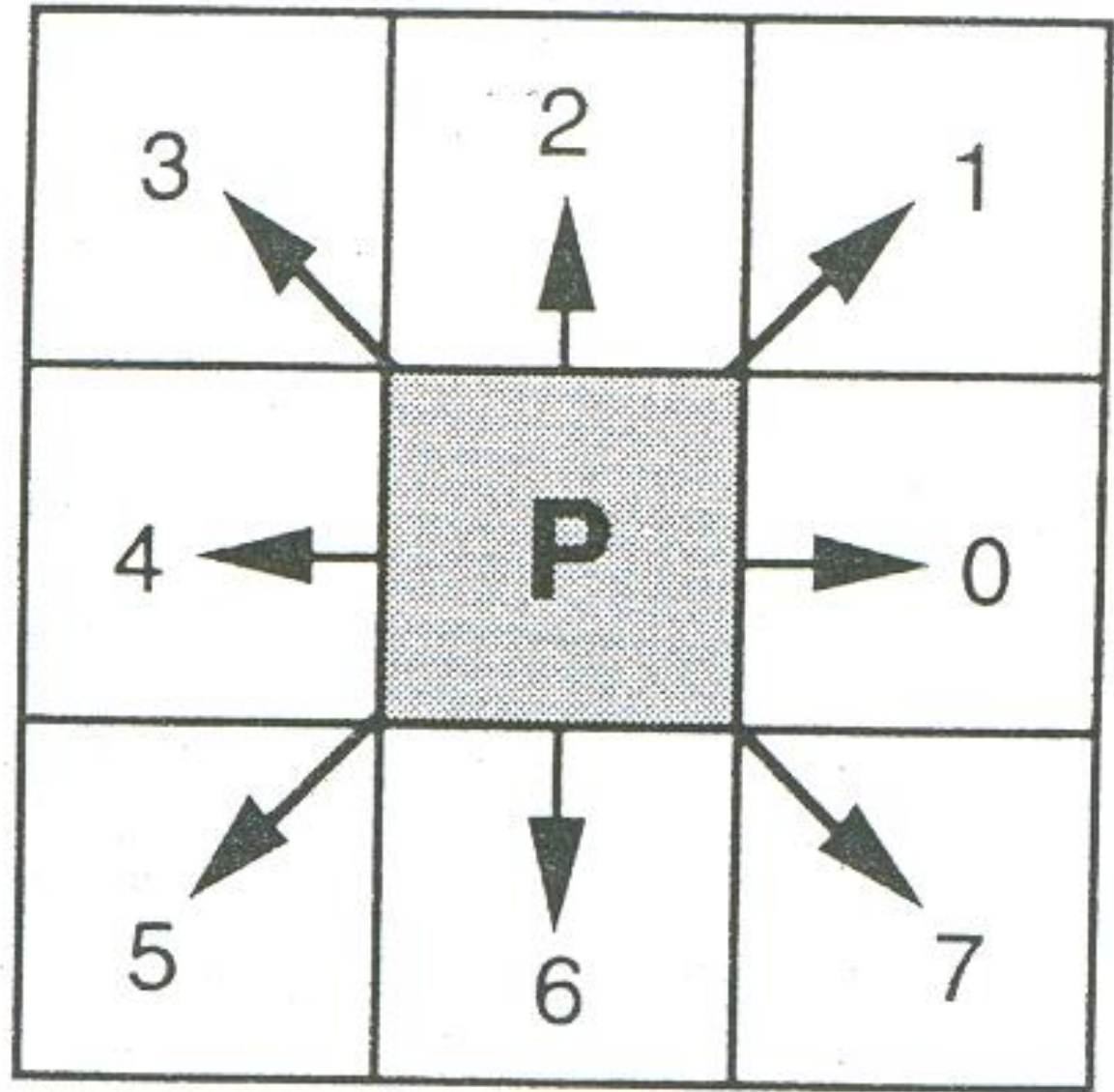
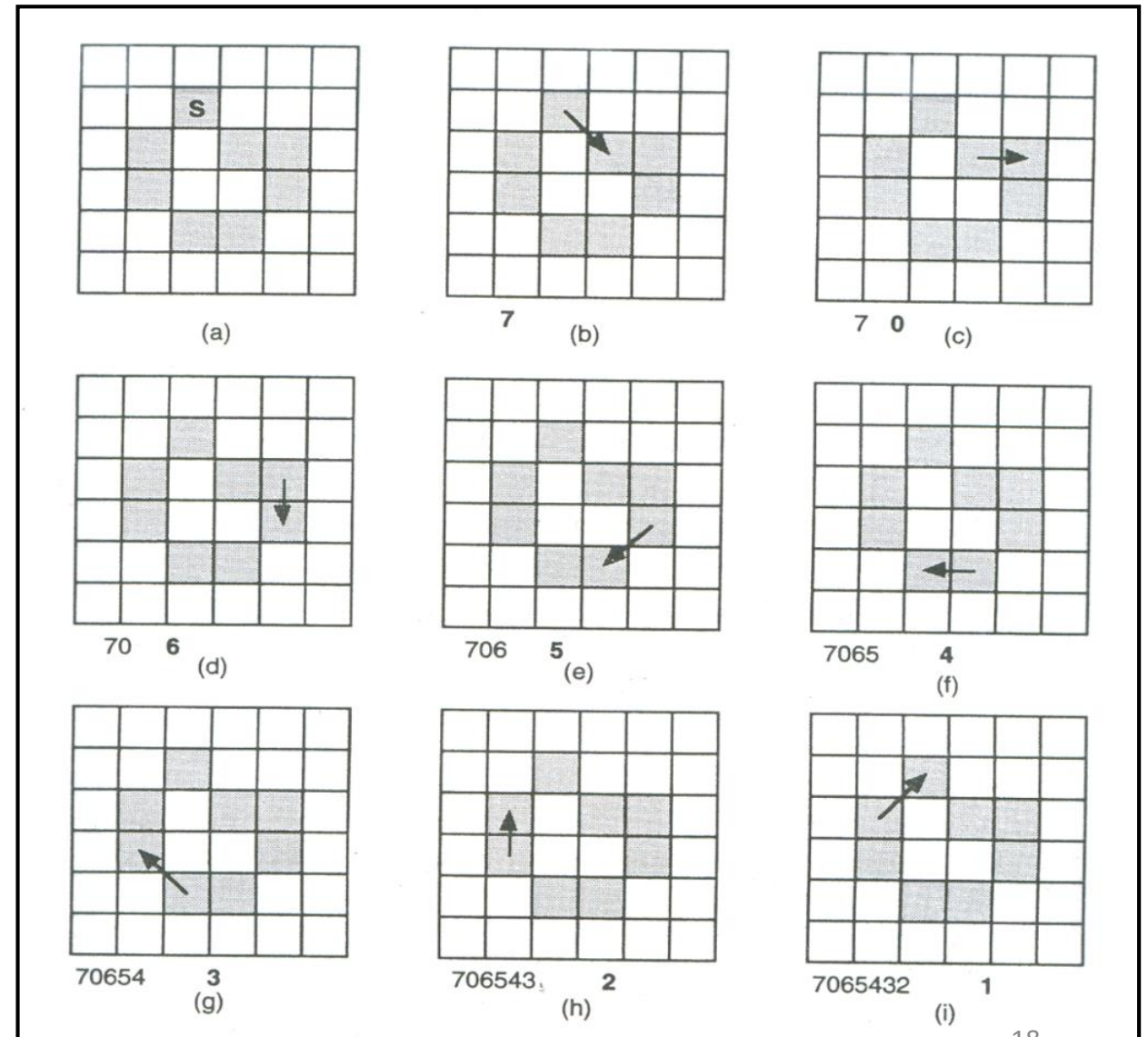
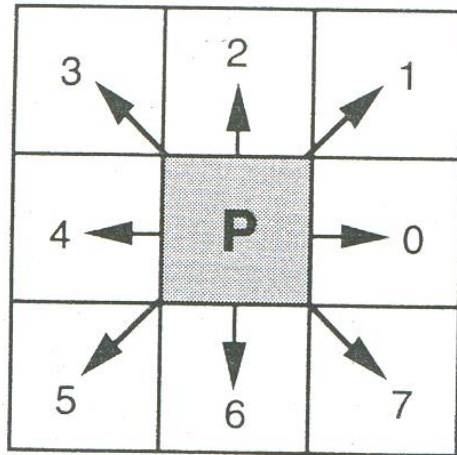


Image Codes

- Technique



Example: Chain Code



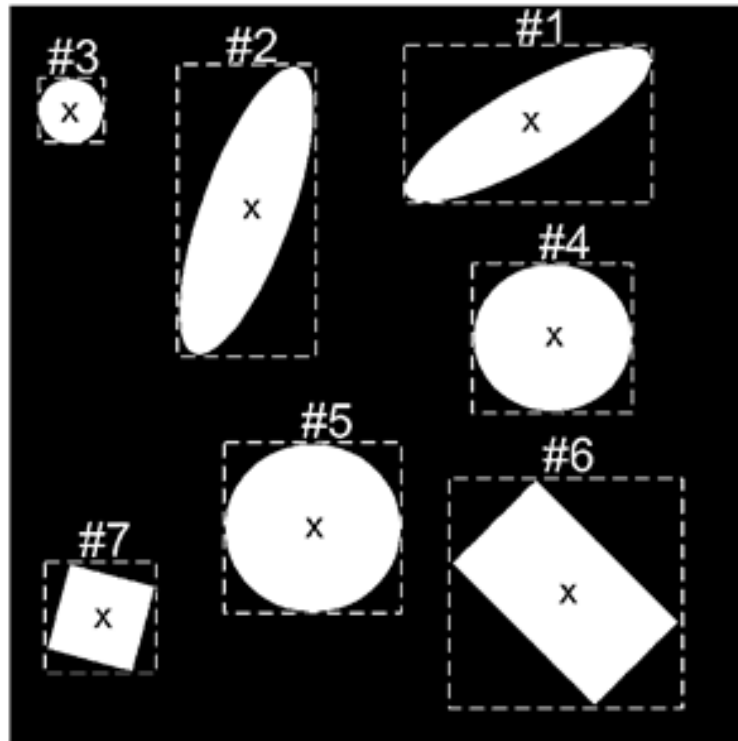
Chain Code properties

- Even codes {0, 2, 4, 6} correspond to horizontal and vertical directions; odd codes {1, 3, 5, 7} correspond to the diagonal directions.
- Each code can be considered as the angular direction, in multiples of 45° , that we must move to go from one contour pixel to the next.
- The absolute coordinates $[m,n]$ of the first contour pixel (e.g. top, leftmost) together with the chain code of the contour represent a complete description of the discrete region contour.
- When there is a change between two consecutive chain codes, then the contour has changed direction. This point is defined as a *corner*.

Boundary-based features

Perimeter	The length of the traced outline
Area	The area enclosed
$X_{\max}$ $X_{\min}$ $Y_{\max}$ $Y_{\min}$	The maximum and minimum co-ordinates reached by the outline
X_{centroid} Y_{centroid}	The X and Y coordinates of the centre of the area enclosed
Shape factor	$\text{Perimeter}^2 / \text{Area}$
Chain count	The number of elements in the chain.

Shape Factor (Circularity)



(a)

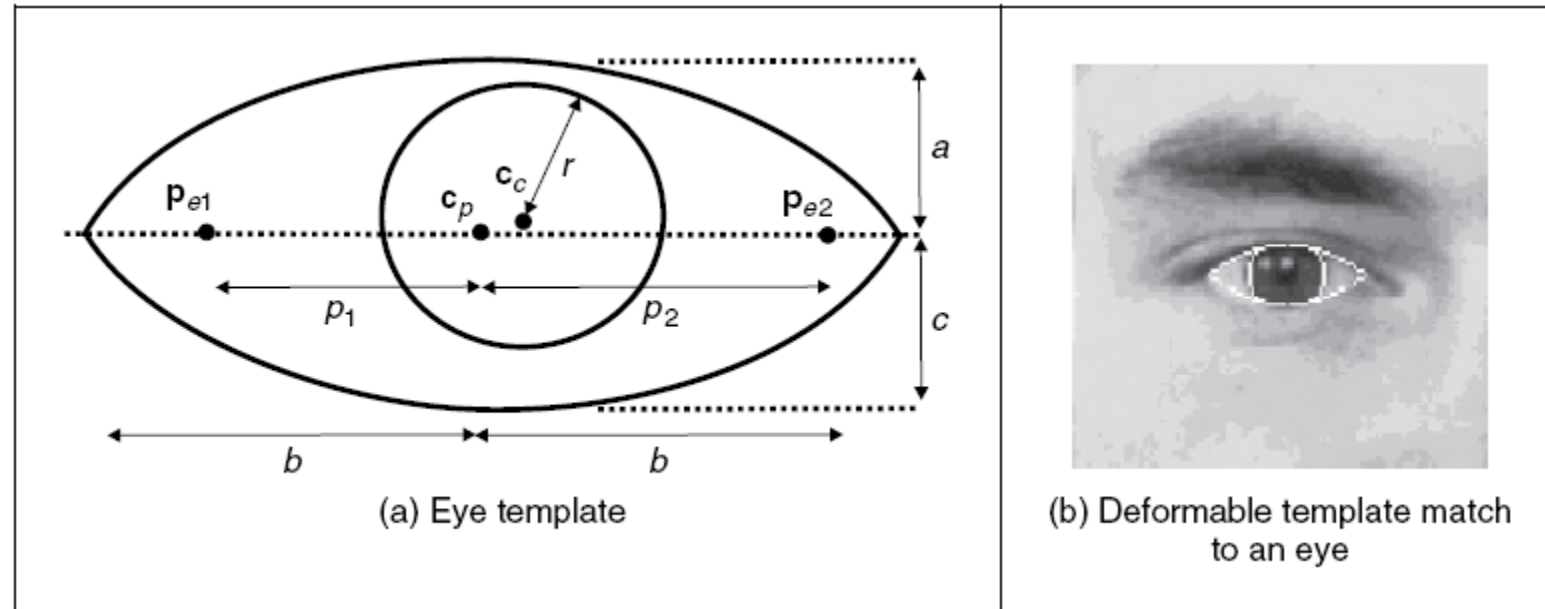
BLOB number	Circularity	Area (pixels)
1	0.31	6561
2	0.40	6544
3	0.98	890
4	0.97	6607
5	0.99	6730
6	0.52	6611
7	0.75	2073

(b)

Active Shape Extraction

Deformable templates	<i>Template matching for deformable shapes. Defining a way to analyse the best match.</i>	Energy maximization, computational considerations, optimization.
Active contours and snakes	<i>Finding shapes by evolving contours. Discrete and continuous formulations. Operational considerations and new active contour approaches.</i>	Energy minimization for curve evolution. Greedy algorithm. Kass snake. Parameterization; initialization and performance. Gradient vector field and level set approaches.
Shape skeletonization	<i>Notions of distance, skeletons and symmetry and its measurement. Application of symmetry detection by evidence gathering. Performance factors.</i>	Distance transform and shape skeleton. Discrete symmetry operator. Accumulating evidence of symmetrical point arrangements. Performance: speed and noise.
Active shape models	<i>Expressing shape variation by statistics. Capturing shape variation within feature extraction.</i>	Active shape model. Active appearance model. Principal components analysis.

Deformable templates



$$\{\mathbf{c}_p, a, b, c, \mathbf{c}_c, r\} = \max \left(\sum_{x,y \in \text{circle.perimeter, parabolas.perimeter}} \mathbf{E}_{x,y} \right)$$

Shape Extraction-Snake Energy Minimization



(a) Initial contour



(b) After the first iteration



(c) After four iterations



(d) After seven iterations

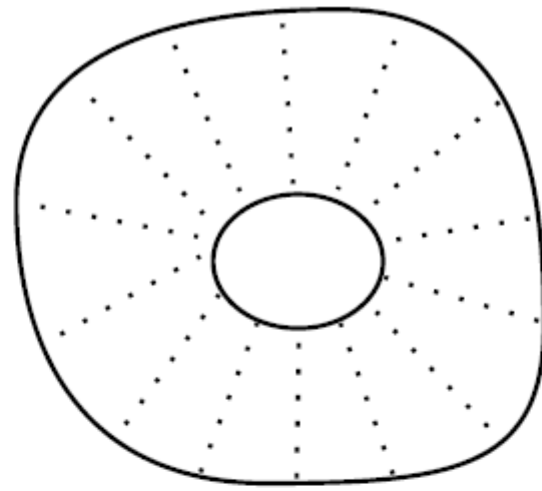
$$E_{\text{snake}} = \int_{s=0}^1 E_{\text{int}}(\mathbf{v}(s)) + E_{\text{image}}(\mathbf{v}(s)) + E_{\text{con}}(\mathbf{v}(s)) ds$$

Shape Extraction-Snake Energy Minimization

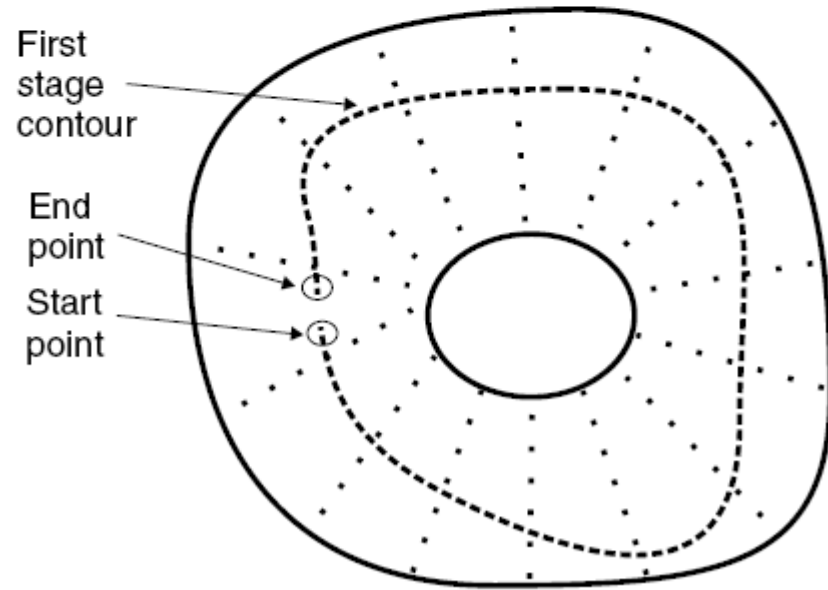
$$E_{\text{snake}} = \int_{s=0}^1 E_{\text{int}}(\mathbf{v}(s)) + E_{\text{image}}(\mathbf{v}(s)) + E_{\text{con}}(\mathbf{v}(s)) ds$$

In this equation, the internal energy, E_{int} , controls the natural behaviour of the snake and hence the arrangement of the snake points; the image energy, E_{image} , attracts the snake to chosen low-level features (such as edge points); and the constraint energy, E_{con} , allows higher level information to control the snake's evolution. The aim of the snake is to evolve by *minimizing*

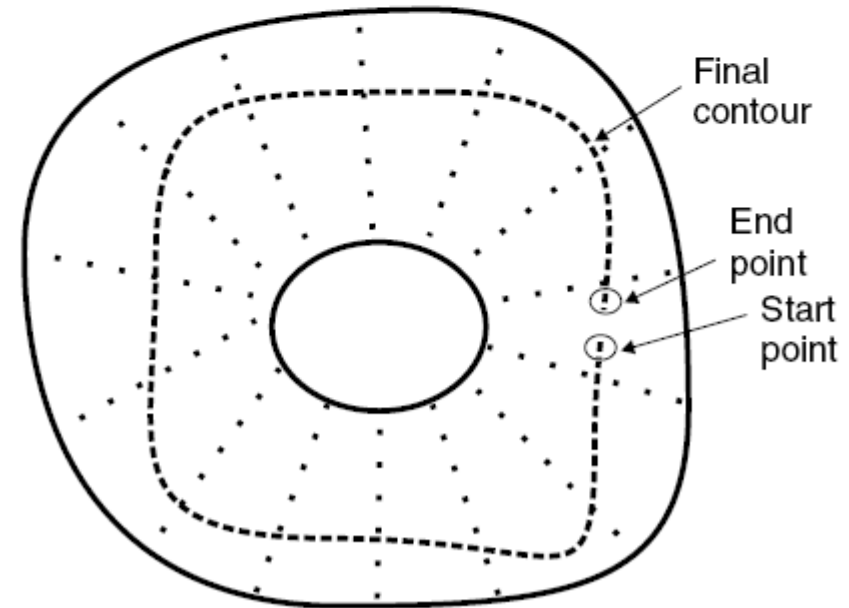
$$\frac{dE_{\text{snake}}}{d\mathbf{v}(s)} = 0$$



(a) Search space



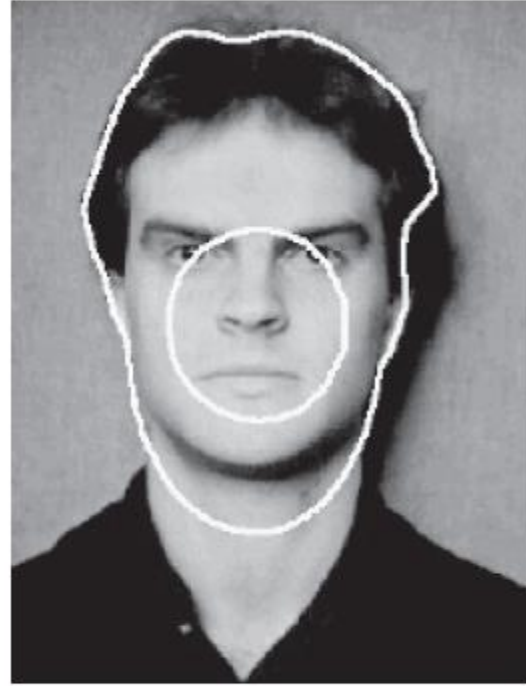
(b) First stage open contour



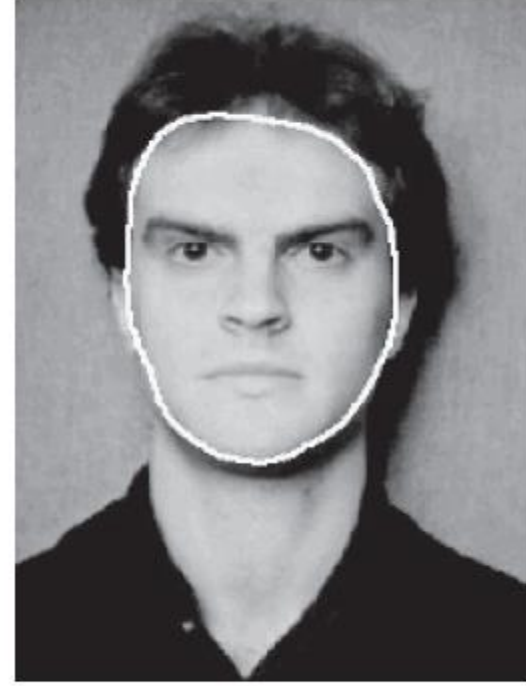
(c) Second stage open contour



(a) Outer boundary initialization



(b) Outer and inner contours



(c) Final face boundary

Geometric Active Contours



(a) Initialization



(b) Iteration 1



(c) Continuing...



(d) Continuing...

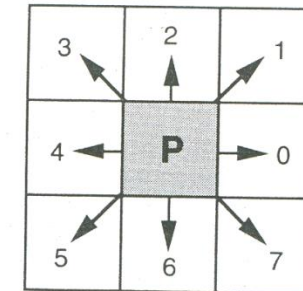
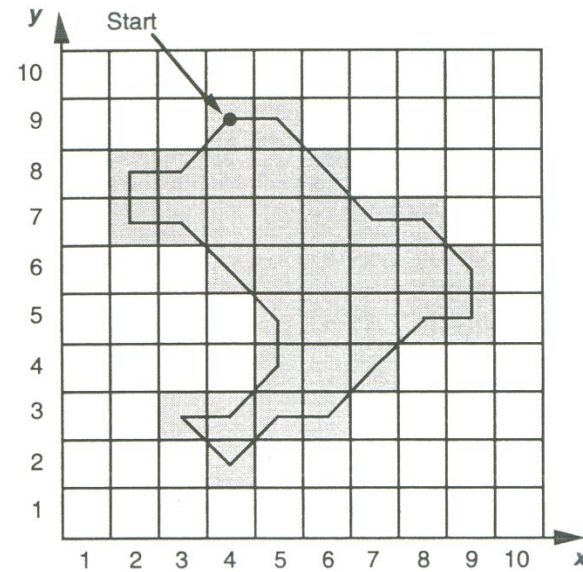


(e) Continuing...



(f) Final result

Example 2: Chain Coding



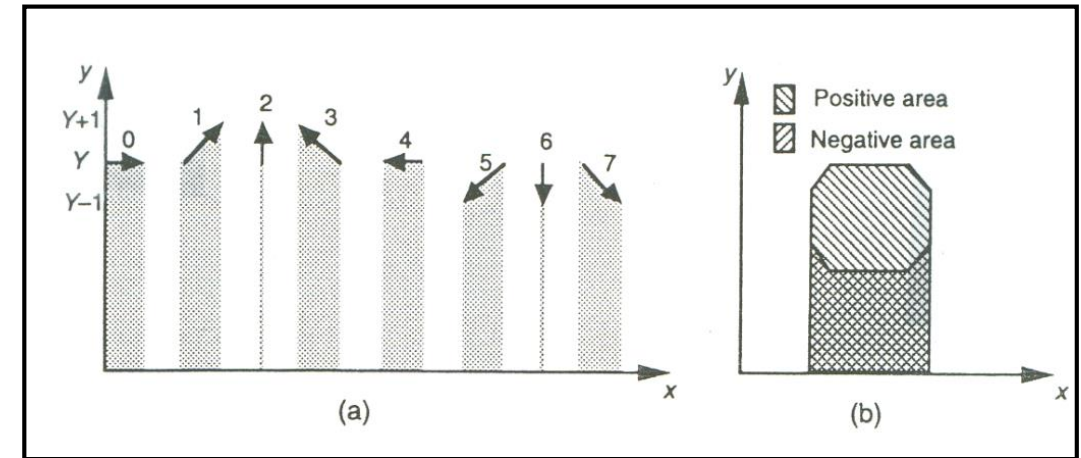
- 077076455453012334201

Calculation of Perimeter

$$\begin{aligned} P &= \text{Even count} + \sqrt{2} \times (\text{Odd count}) \\ &= 10 + \sqrt{2} \times 11 = 25.56 \text{ units} \end{aligned}$$

Calculation of Area using Chain Code

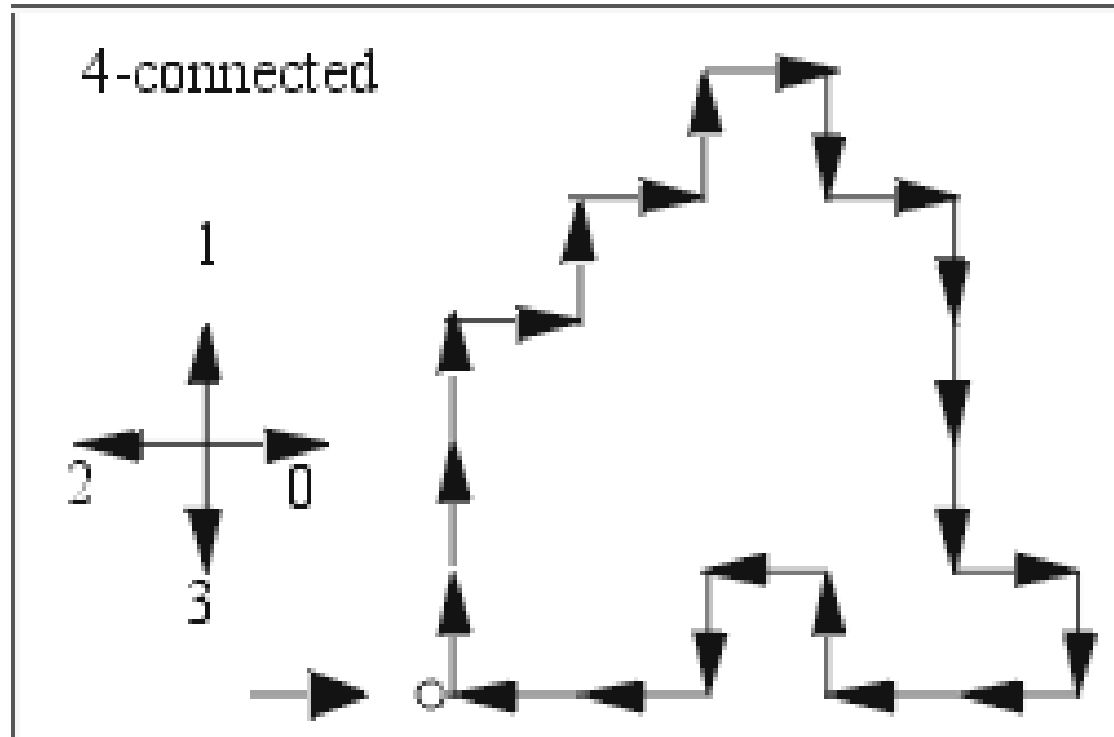
Chain code vector 0	Area: = Area + Y
Chain code vector 1	Area: = Area + (Y+0.5)
Chain code vector 2	Area: = Area
Chain code vector 3	Area: = Area - (Y+0.5)
Chain code vector 4	Area: = Area - Y
Chain code vector 5	Area: = Area - (Y-0.5)
Chain code vector 6	Area: = Area
Chain code vector 7	Area: = Area + (Y-0.5)



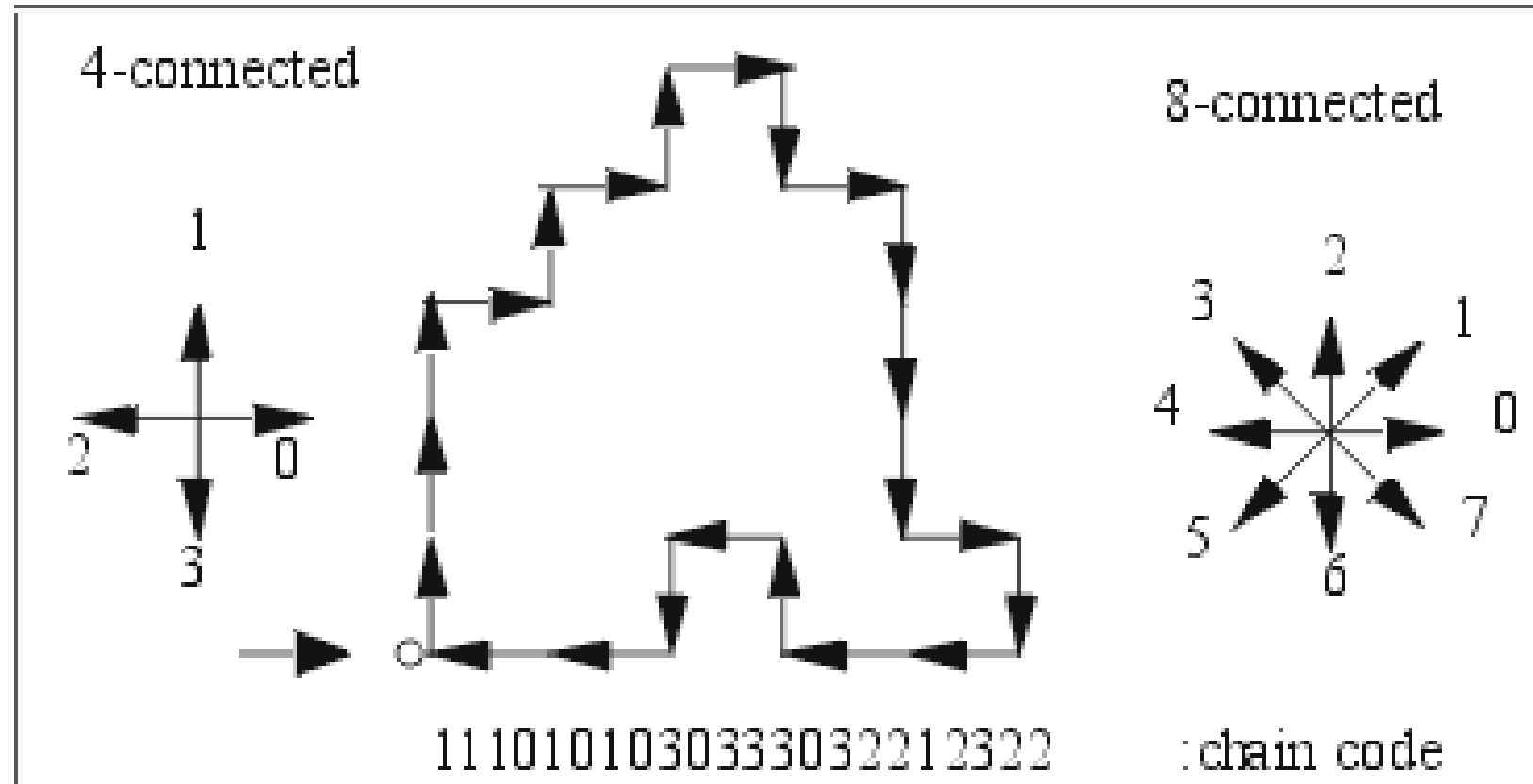
077076455453012334201

Vector Code:	0	7	7	0	7	6	4	5	etc.
Ordinate:	9	9	8	7	7	6	5	5	etc.
Area:	+9	+8.5	+7.5	+7	+6.5	+0	-5	-4.5	etc.
Cumulative area:	9	17.5	25	32	38.5	38.5	33.5	29	etc.

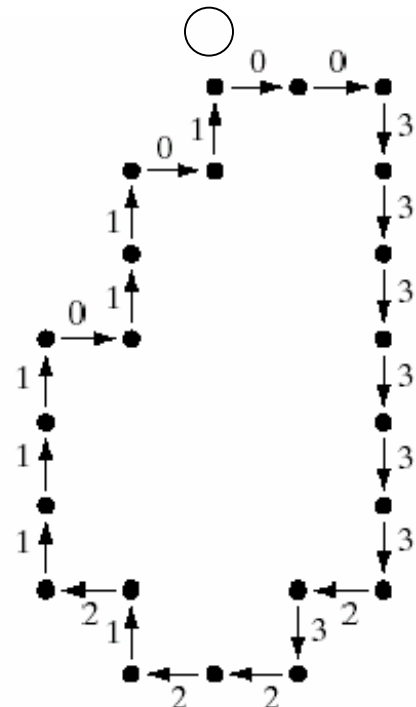
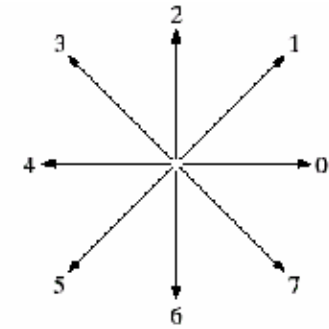
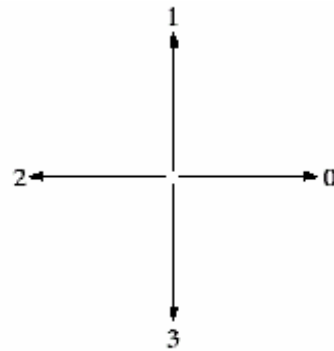
Freeman Chain Code



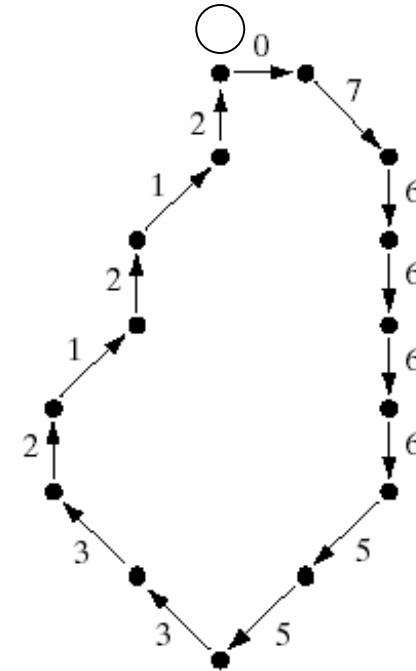
Freeman Chain Code



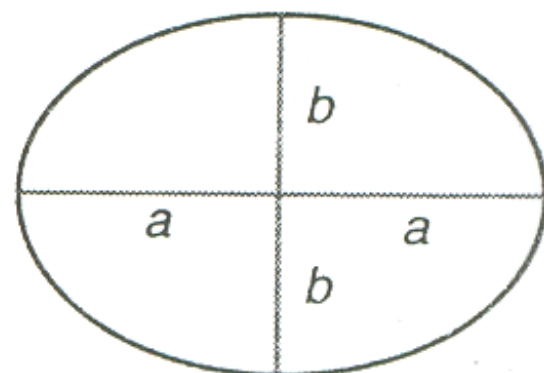
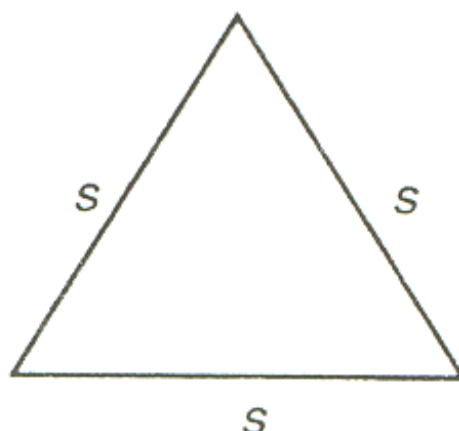
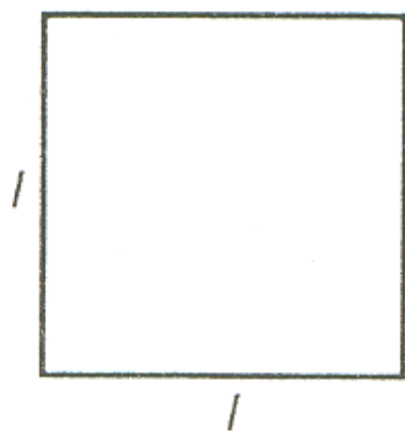
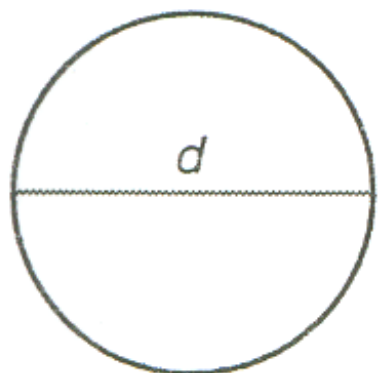
Chain Codes



4-directional chain code:
003333323221211101101



8-directional chain code:
076666553321212



$$P = \pi d$$

$$A = \pi d^2/4$$

$$S = 4\pi = 12.6$$

$$P = 4l$$

$$A = l^2$$

$$S = 16$$

$$P = 3s$$

$$A = (s^2\sqrt{3})/4$$

$$S = 36/\sqrt{3} = 20.8$$

$$P = 2\pi\sqrt{\{(a^2 + b^2)/2\}}$$

$$A = \pi ab$$

$$S = \{2\pi(a^2 + b^2)\}/ab$$

Binary Object Features – Aspect Ratio

- The aspect ratio (also called elongation or eccentricity) is defined by the ratio of the bounding box of an object.
- This can be found by scanning the image and finding the minimum and maximum values on the row and column where the object lies.

Binary Object Features – Aspect Ratio

- The equation for aspect ratio is as follows:

$$\frac{c_{\max} - c_{\min} + 1}{r_{\max} - r_{\min} + 1}$$

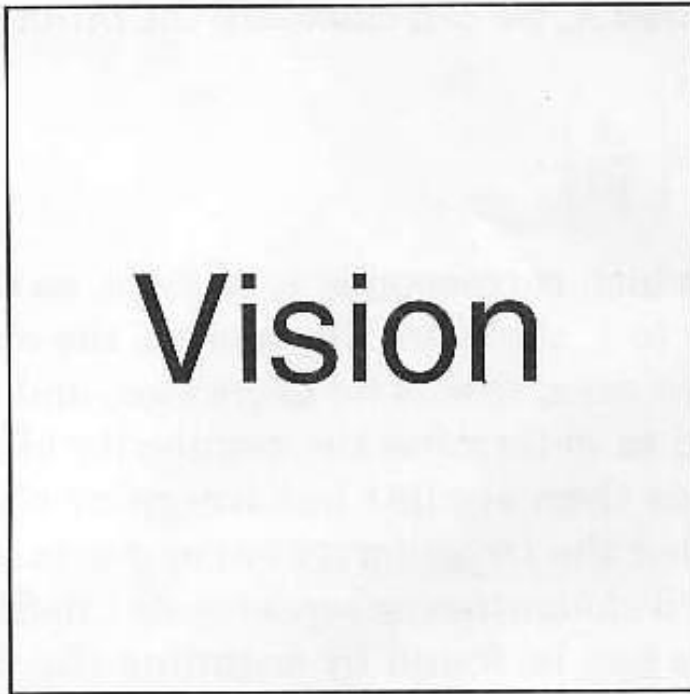
- It reveals how the object spread in both vertical and horizontal direction.
- High aspect ratio indicates the object spread more towards horizontal direction.

Binary Object Features – Euler Number

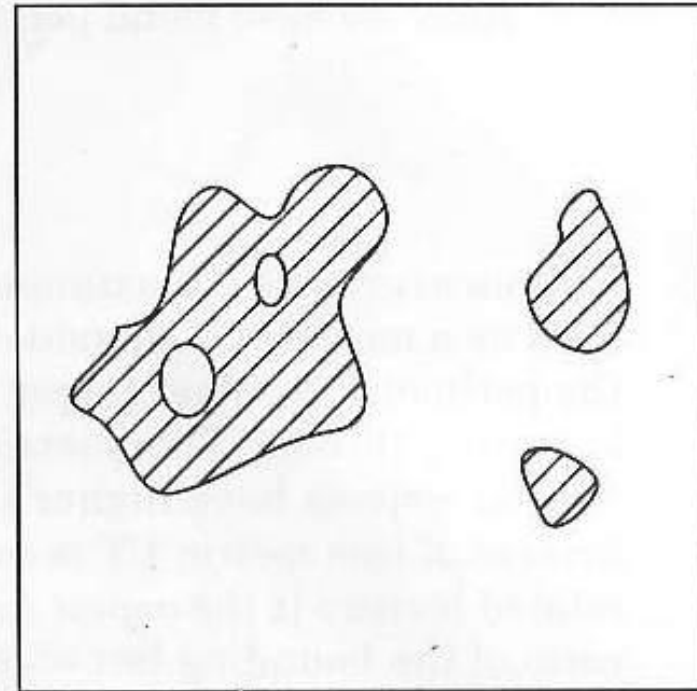
- Euler number is defined as the difference between the number of objects and the number of holes.
 - Euler number = num of object – number of holes
- In the case of a single object, the Euler number indicates how many closed curves (holes) the object contains.

Binary Object Features – Euler Number

- Euler number can be used in tasks such as optical character recognition (OCR).

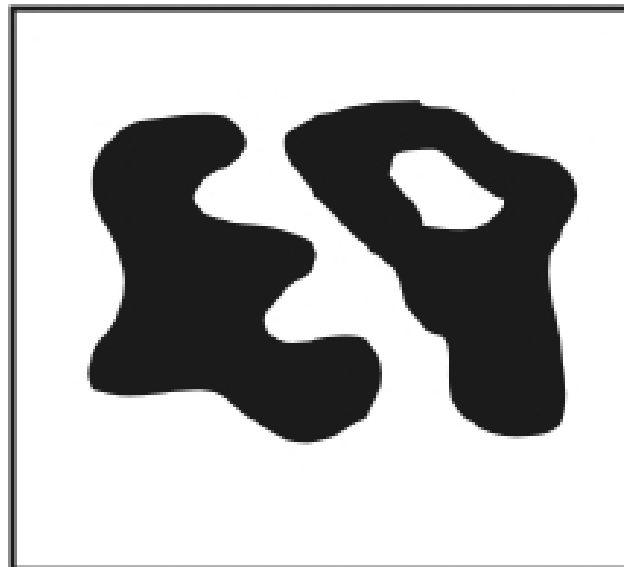


a. This image has eight objects and one hole, so its Euler number is $8 - 1 = 7$. The letter *V* has Euler number of 1, $i = 2$, $s = 1$, $o = 0$, and $n = 1$.

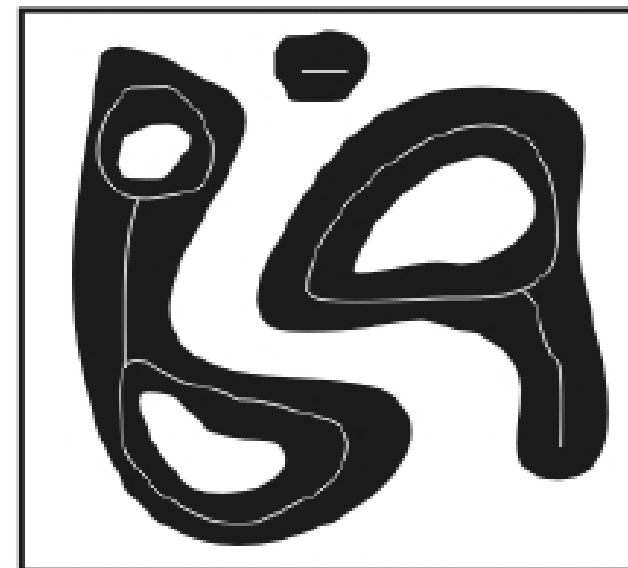
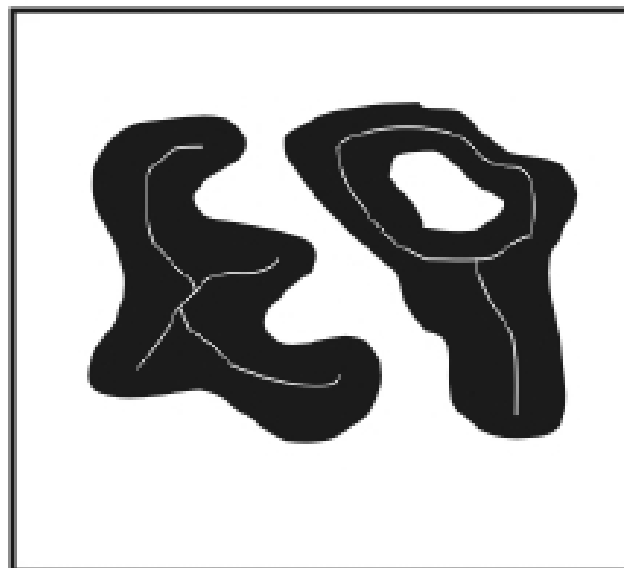


b. This image has three objects and two holes, so the Euler number is $3 - 2 = 1$.

Euler Number



(a)



(b)

Application	How Euler Number Helps
Character recognition	Differentiates characters based on holes (e.g., 'A' vs. 'H', '0' vs. '8')
Object counting	Identifies the number of objects and their topological complexity
Shape analysis	Helps classify objects by counting holes or enclosed spaces
Noise detection	Detects unwanted small holes or extra components in segmented images
Medical imaging	Analyzes cell or tissue structures based on their connectivity and cavity counts

Binary Object Features – Projection

- The projection of a binary object, may provide useful information related to object's shape.
- It can be found by summing all the pixels along the rows or columns.
 - Summing the rows give horizontal projection.
 - Summing the columns give the vertical projection.

Binary Object Features – Projection

- We can defined the horizontal projection $h_i(r)$ and vertical projection $v_i(c)$ as:

$$h_i(r) = \sum_{c=0}^{width-1} I_i(r, c)$$

$$v_i(c) = \sum_{r=0}^{height-1} I_i(r, c)$$

- An example of projections is shown in the next slide:

Binary Object Features – Projection

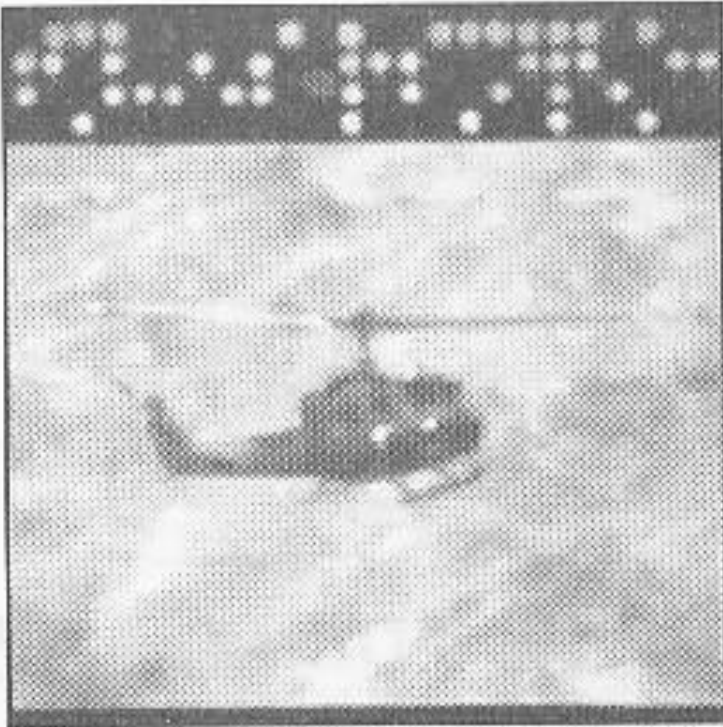
									$h(r)$
									↓
$c \rightarrow$									
r ↓	0	0	0	0	1	0	0	0	1
	0	0	0	0	1	1	0	0	2
	0	0	1	1	1	1	1	0	5
	0	1	1	1	1	1	1	0	6
	0	0	0	0	0	1	1	0	2
	0	0	0	0	0	0	0	0	0
$v(c) \rightarrow$	0	1	2	2	4	4	3	0	0

To find the projections, we sum the number of 1s in the rows and columns.

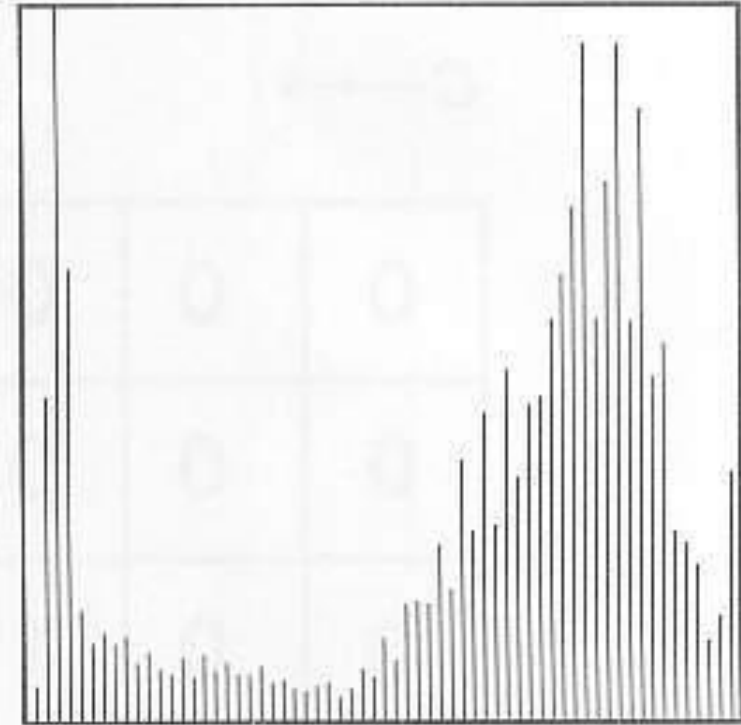
Histogram Features

- The histogram of an image is a plot of the gray-level values versus the number of pixels at that value.
- The shape of the histogram provides us with information about the nature of the image.
 - The characteristics of the histogram has close relationship with characteristic of image such as brightness and contrast.

Histogram Features

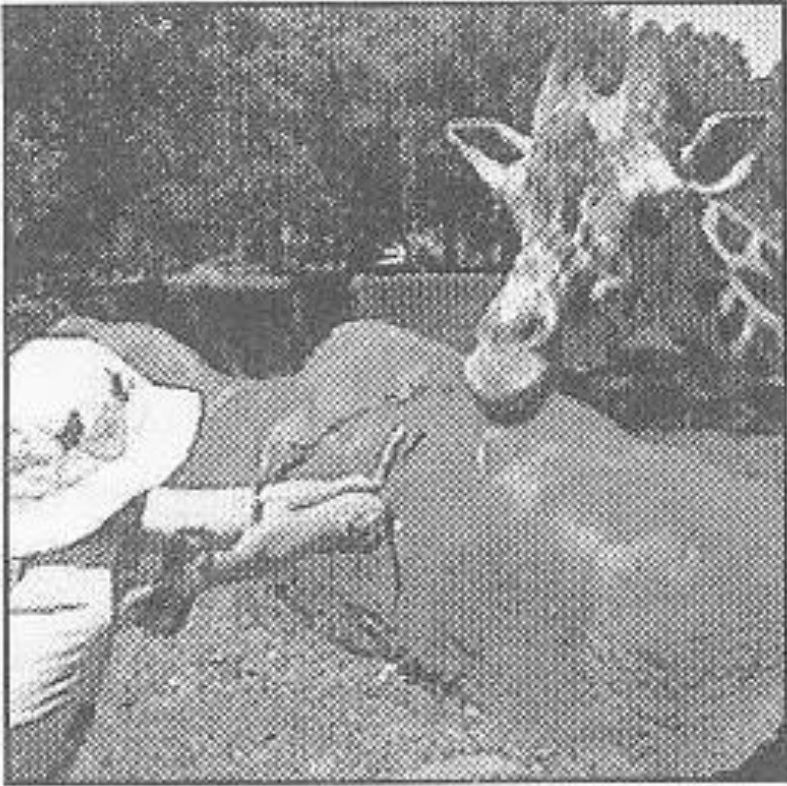


a. Object in contrast with background.

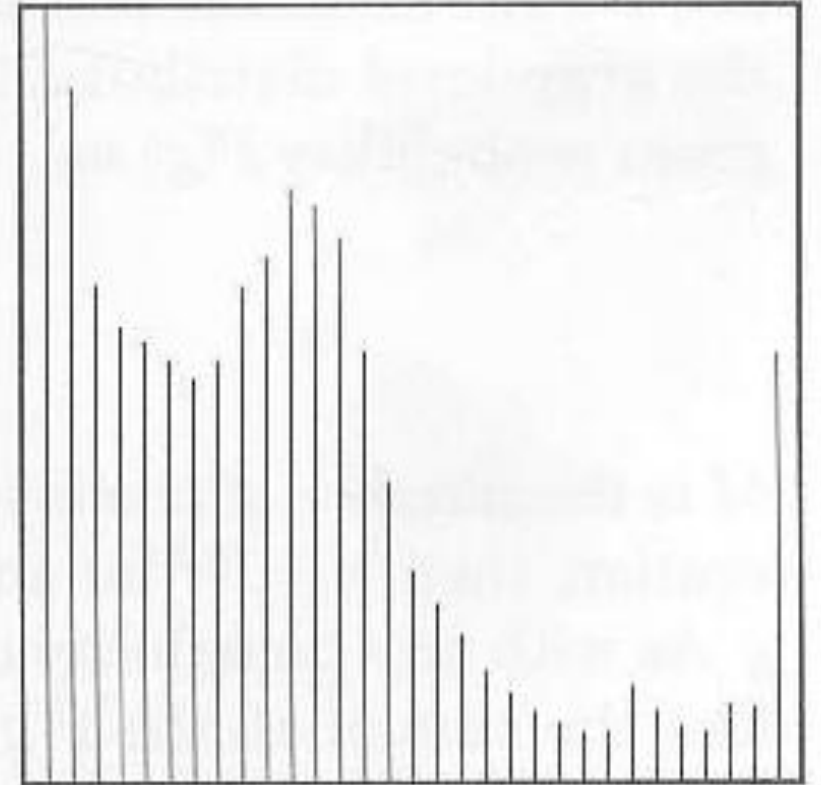


b. Histogram of (a) shows bimodal shape.

Histogram Features



e. High-contrast image.

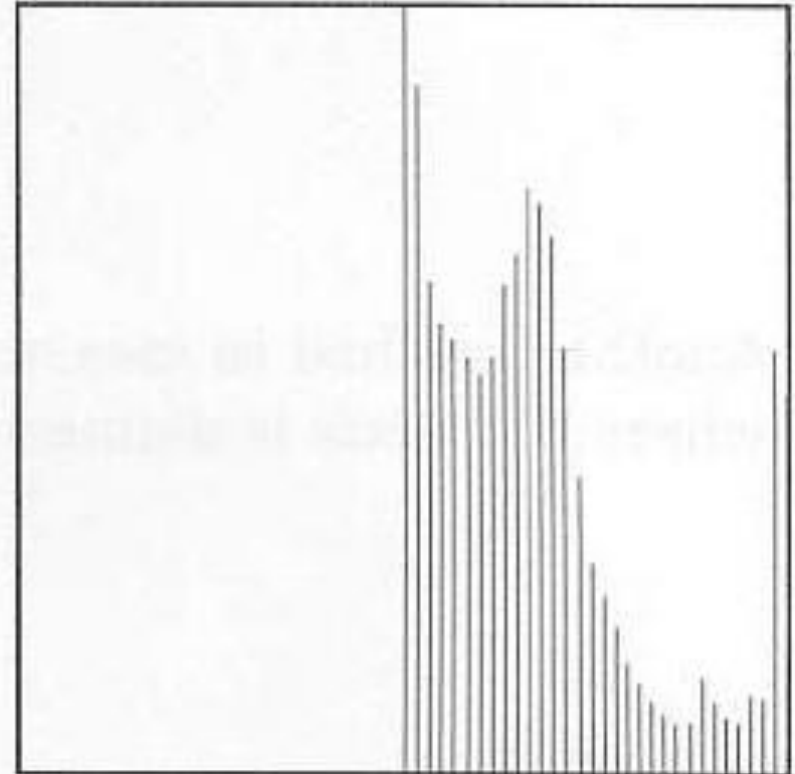


f. Histogram of (e) appears spread out.

Histogram Features



g. Bright image.

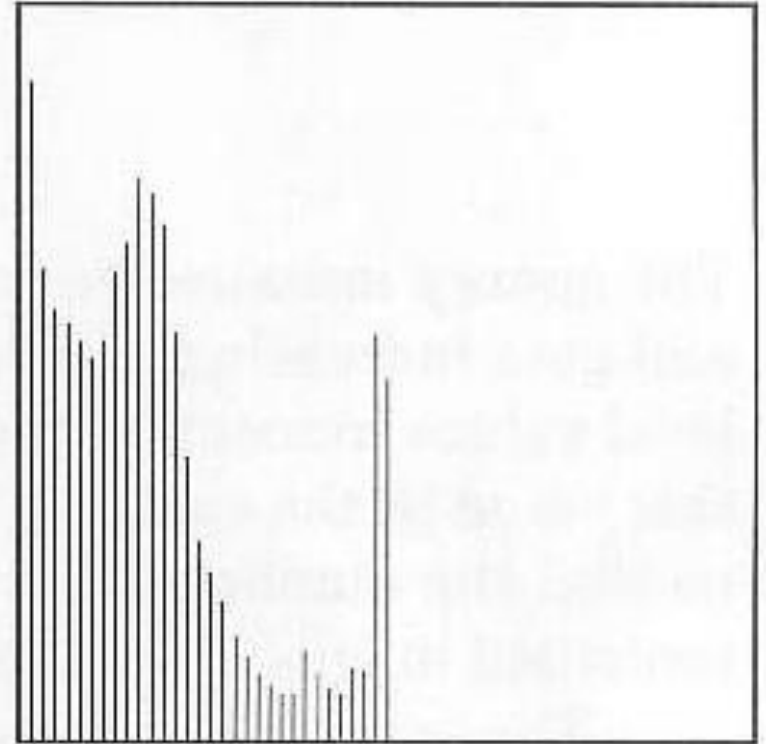


h. Histogram of (g) appears shifted to the right.

Histogram Features



i. Dark image.



j. Histogram of (i) appears shifted to the left.

Histogram Features

- The features based on the first-order histogram probability are
 - Mean
 - Standard deviation
 - Skew
 - Energy
 - Entropy.

Histogram Features – Mean

- The mean is the average value, so it tells us something about the general brightness of the image.
 - A bright image has a high mean.
 - A dark image has a low mean.
- The mean can be defined as follows:

$$Mean = \bar{g} = \sum_{g=0}^{L-1} gP(g) = \sum_{r=0}^{height-1} \sum_{c=0}^{width-1} \frac{I(r, c)}{M}$$

Histogram Features – Standard Deviation

- The standard deviation, which is also known as the square root of the variance, tells something about the contrast.
- It describes the spread in the data.
 - Image with high contrast should have a high standard deviation.
- The standard deviation is defined as follows:

$$\sigma_g = \sqrt{\sum_{g=0}^{L-1} (g - \bar{g})^2 P(g)}$$

Histogram Features – Skew

- The skew measures the asymmetry (unbalance) about the mean in the gray-level distribution.
- Image with bimodal histogram distribution (object in contrast background) should have high standard deviation but low skew distribution (one peak at each side of mean).

Histogram Features – Skew

- Skew can be defined in two ways:

$$SKEW = \frac{1}{\sigma_g^3} \sum_{g=0}^{L-1} (g - \bar{g})^3 P(g)$$

$$SKEW' = \frac{\bar{g} - \text{mod}}{\sigma_g}$$

- In the second method, the *mod* is defined as the peak, or highest value.

Histogram Features – Energy

- The energy measure tells us something about how gray levels are distributed.
- The equation for energy is as follows:

$$ENERGY = \sum_{g=0}^{L-1} [P(g)]^2$$

Histogram Features – Energy

- The energy measure has a value of 1 for an image with a constant value.
- This value gets smaller as the pixel values are distributed across more gray level values.
- A high energy means the number of gray levels in the image is few.
 - Therefore it is easier to compress the image data.

Histogram Features – Entropy

- Entropy measures how many bits do we need to code the image data.
- The equation for entropy is as follows:

$$ENTROPY = - \sum_{g=0}^{L-1} P(g) \log_2 [P(g)]$$

- As the pixel values are distributed among more gray levels, the entropy increases.

Spectral Images

- The primary metric for spectral features (frequency-domain-based features) is power.
- Power is defined as the magnitude of the spectral component squared.

$$POWER = |T(u, v)|^2$$

- Spectral features are useful when classifying images based on textures.
 - Done by looking for peaks in the power spectrum.

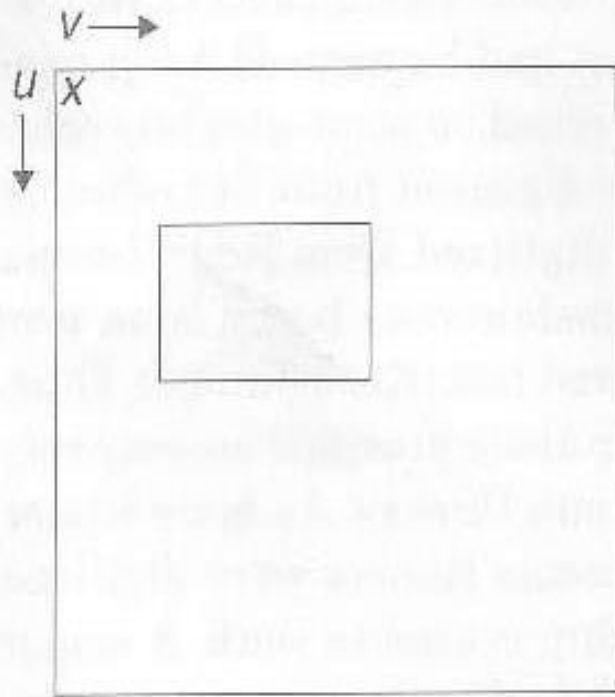
Spectral Images

- It is typical to look at power in various regions, and these regions can be defined as rings, sectors or boxes.
- We can then measure the power in a region of interest by summing the power over the range of frequencies of interest.

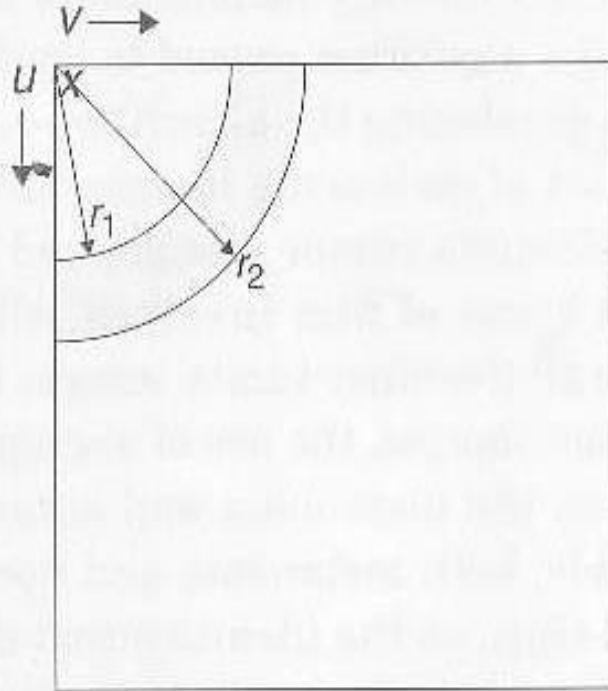
$$\textit{Spectral RegionPower} = \sum_{u \in \textit{REGION}} \sum_{v \in \textit{REGION}} |T(u, v)|^2$$

Spectral Images

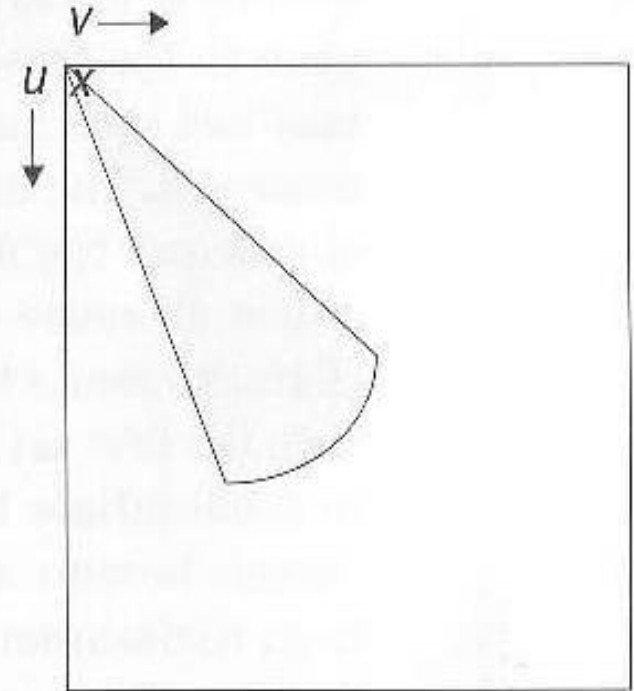
Cosine and Walsh-Hadamard Transform Symmetry
 $x = \text{origin}$



d. Box symmetry.



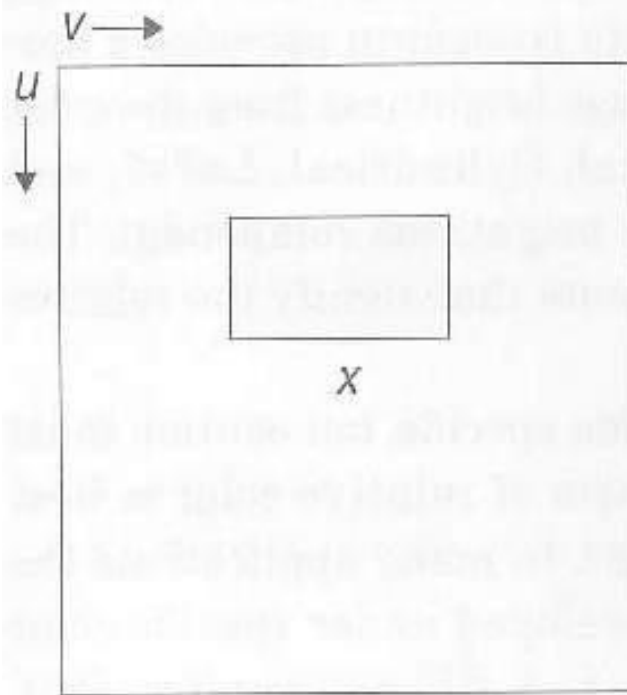
e. Ring symmetry.



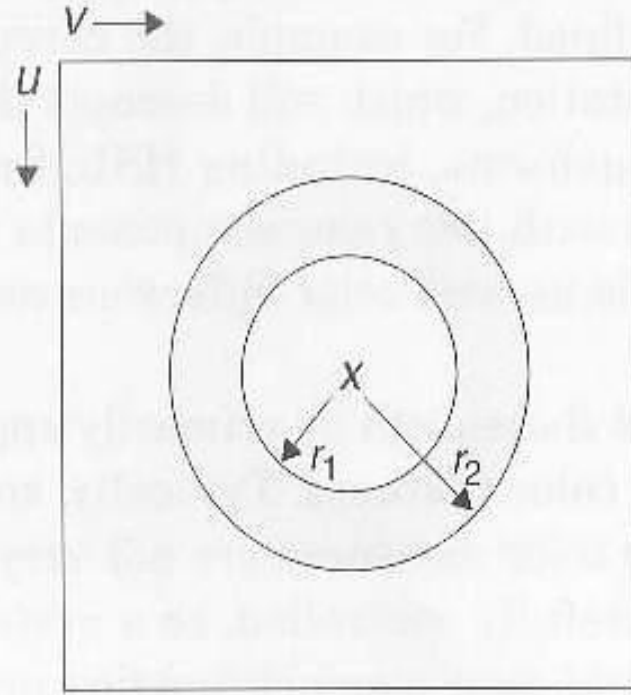
f. Sector symmetry.

Spectral Images

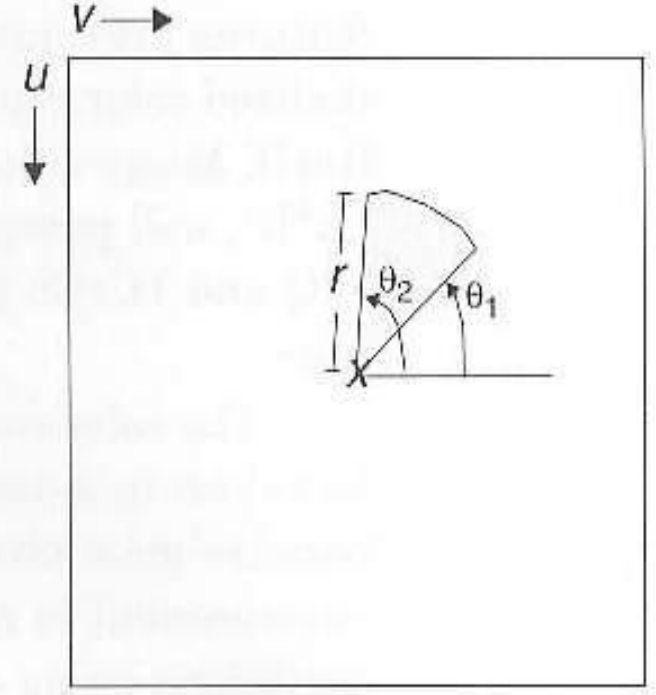
Fourier Transform Symmetry
 $x = \text{origin}$



a. Box is defined by limits on u and v .



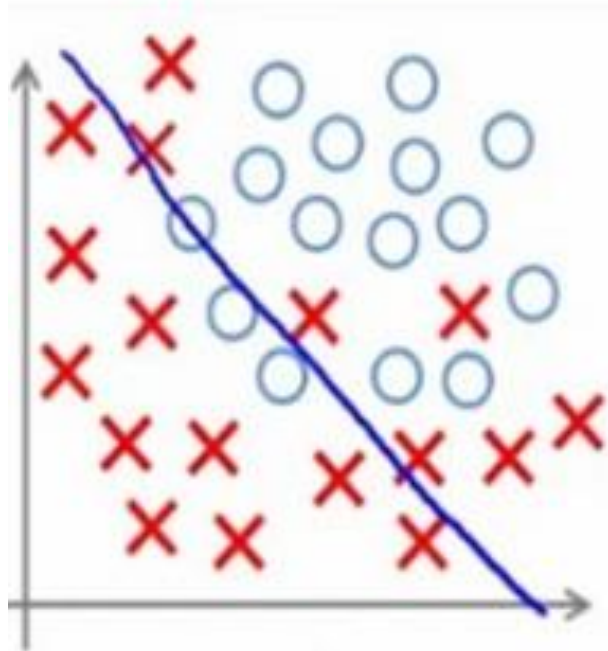
b. Ring is defined by limits on the radii from origin x .



c. Sector is defined by radius r and angles θ_1 and θ_2 .

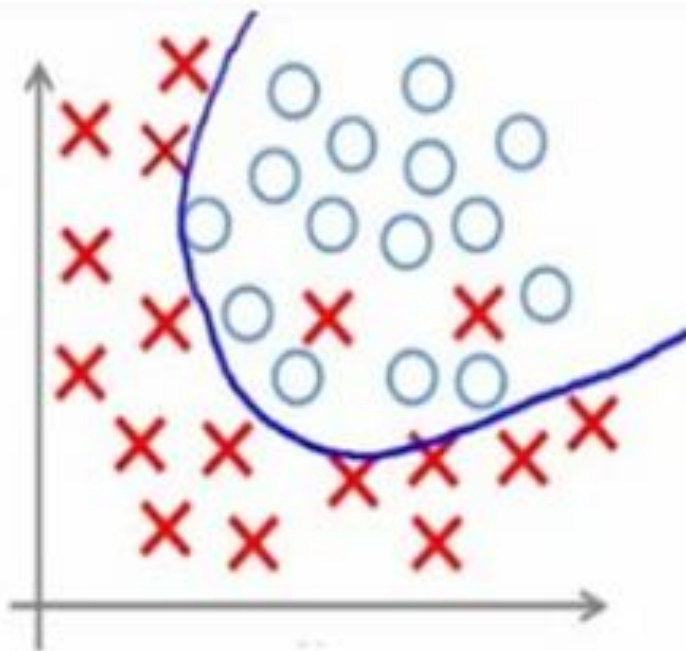
Spectral Images

- The ring measure can be used to find texture:
 - High power in small radii corresponds to smooth textures.
 - High power in large radii corresponds to coarse texture.
- The sector power measure can be used to find lines or edges in a given direction, but the results are size invariant.

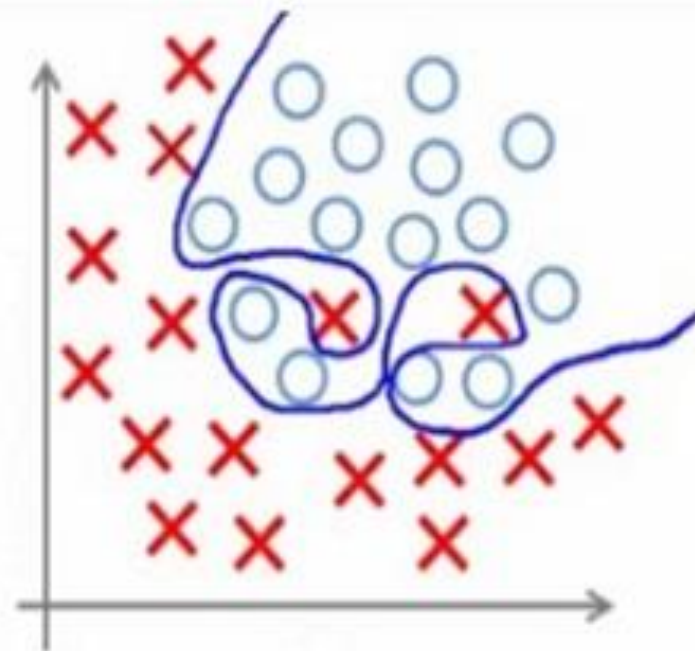


Under-fitting

(too simple to
explain the
variance)



Appropriate-fitting



Over-fitting

(forcefitting – too
good to be true)

Machine Learning Approaches

